

SEQUENCE LISTING



GENERAL INFORMATION:

- (i) APPLICANT: White, Theresa C  
Hindle, Christopher D
- (ii) TITLE OF INVENTION: GENETIC CONSTRUCTS AND GENETICALLY  
MODIFIED MICROBES FOR ENHANCED PRODUCTION OF  
BETA-GLUCOSIDASE.
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Fitzpatrick, Cella, Harper & Scinto
  - (B) STREET: 1900 K Street, N.W.
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: USA
  - (F) ZIP: 20006-1110
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: CA 09/037,524
  - (B) FILING DATE: 10-MAR-1998
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Olsen, Warren E
  - (C) REFERENCE/DOCKET NUMBER: 1039.2500
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-530-1010
  - (B) TELEFAX: 202-530-1055

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 72 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Trichoderma reesei*

**(ix) FEATURE:**

(A) NAME/KEY: sig\_peptide

(B) LOCATION: 1..51

(D) OTHER INFORMATION: /function= "cbhl secretion signal"

(ix) **FEATURE:**

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 61.72

(D) OTHER INFORMATION: /function= "mature  
beta-glucosidase"

(ix) **FEATURE:**

(A) NAME/KEY: CDS

(B) LOCATION: 1..72

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG TAT CGG AAG TTG GCC GTC ATC TCG GCC TTC TTG GCC ACA GCT CGT 48  
Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg  
-20 -15 -10 -5

GCT CAG TCG GCA GTT GTA CCT CCT 72  
Ala Gln Ser Ala Val Val Pro Pro  
1

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Tyr Arg Lys Leu Ala Val Ile Ser Ala Phe Leu Ala Thr Ala Arg  
-20 -15 -10 -5

Ala Gln Ser Ala Val Val Pro Pro  
1

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Trichoderma reesei

(ix) FEATURE:

(A) NAME/KEY: sig\_peptide

(B) LOCATION: 1..99

(D) OTHER INFORMATION: /function= "xln 2 signal peptide"

(ix) FEATURE:

(A) NAME/KEY: mat\_peptide

(B) LOCATION: 109..120

(D) OTHER INFORMATION: /function= "mature  
beta-glucosidase"

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATG GTC TCC TTC ACC TCC CTC CTC GCC GGC GTC GCC GCC ATC TCG GGC	48
Met Val Ser Phe Thr Ser Leu Leu Ala Gly Val Ala Ala Ile Ser Gly	
-36 -35 -30 -25	
GTC TTG GCC GCT CCC GCC GCC GAG GTC GAA TCC GTG GCT GTG GAG AAG	96
Val Leu Ala Ala Pro Ala Ala Glu Val Glu Ser Val Ala Val Glu Lys	
-20 -15 -10 -5	
CGC CAG GCT AGA GTT GTA CCT CCT	120
Arg Gln Ala Arg Val Val Pro Pro	
1	

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Val Ser Phe Thr Ser Leu Leu Ala Gly Val Ala Ala Ile Ser Gly
-36 -35 -30 -25
Val Leu Ala Ala Pro Ala Ala Glu Val Glu Ser Val Ala Val Glu Lys

-20

-15

-10

-5

Arg Gln Ala Arg Val Val Pro Pro  
1